Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-61-AC2, 2020

Geoscientific Model Development Discussions

Interactive comment on "Model-driven optimization of coastal sea observatories through data assimilation in a finite element hydrodynamic model (SHYFEM v.7_5_65)" *by* Christian Ferrarin et al.

Christian Ferrarin et al.

c.ferrarin@ismar.cnr.it

Received and published: 11 November 2020

Dear Reviewer,

We would like to thank you very much for your tireless efforts in reviewing the manuscript and for your valuables comments, which will certainly improve our work.

The original Reviewer's comments and suggestions are shown in regular typeface, while our responses are shown in italics.

Sincerely yours, Christian Ferrarin (On behalf of the authors)

R2.1 General comments: The paper analyzes optimization of observational grid via analyzing the impact that assimilation of station data has on the high resolution numerical model of the Venice lagoon. Several modes of assimilation are employed to introduce data into the model. I must say I really like the idea of how DA was used in the paper. The paper is interesting, contains new insight and is well written. The figures are clear. The abstract reflects the contents well.

I recommend publication after minor revision.

Response: We thank the reviewer for the positive comment and we improved the manuscript following all reviewer's suggestions.

R2.2 p3, L76: pa should be p_a ($_a$ denoting subscript)

Response: Corrected.

R2.3 p3, L77: ρ_q should be ρ_w

Response: Corrected.

R2.4 p4, L107: "the mean is:" should probably be "the ensemble mean is:"

Response: Corrected.

R2.5 p7 L185: I am not sure I understand this phrase "... at which degree the observations represent the state variable over the whole system." Can the authors please include a specific description and/or metrics by which this degree was measured?

Response: As explained in section 3.1, for both Data Interpolation and Data Assimilation experiments the metric used to evaluate the representativeness of the method in describing the state variable over the system is the root mean square

error. RMSE is evaluated in the station not considered in the DI or DA computations. The evaluation procedure was repeated for each monitoring station and the results are reported in Table 1.

R2.6 P8 L207: should sigma be a greek letter? Why did you set it to 2 km rather than something else?

Response: We corrected the sigma Greek letter. As specified at the beginning of the Results section, all parameters (σ, τ, cut-off distance for the local analysis) were manually defined through trial and error calibration process and evaluating the goodness-of-fit of the water level RMSE in the DA-Nudging and DA-EnSRF base simulations.

R2.7 Perhaps I missed something but I still do not clearly understand how the boundary condition perturbations were generated. The paper states that 60 perturbations (gaussian, it seems?) were used as OBCs. Do I understand correctly that you used mean(A) as the open boundary conditions and then further added a constant (in space and time) perturbation to each ensemble member, where the amount of each member sea level perturbation was sampled from a gaussian N(mu, sigma)?

Response: We concur with the reviewer that the perturbation terms were not properly described. At each timestep (t), a random vector (r) of N perturbations is computed from a Gaussian distribution (with mean 0 and standard deviation of 30 cm) as:

$$r(t,n) = cos(2\pi r_2(t,n))\sqrt{-2log(r_1(t,n)+\epsilon)} \quad (1)$$

with n the number of the ensemble member ($1, N$), r_1 and r_2 random vectors and ϵ a very small number.

The perturbation vector p at time t is computed using the random vector and the

perturbation vector at the previous time (t_{-1}):

$$p(t,n) = \alpha p(t_{-1},n) + \sqrt{1-\alpha^2} r(t,n) \quad (2)$$

with $\alpha = 1-(t-t_{-1})/\tau$ and τ the decay time (2 days in our case). Then the new perturbation is stored for the next time step. This type of perturbations are classified as red noise.

We modified the manuscript to clarify the methodology adopted in this study. The text now reads: "We used 60 perturbations for the sea-level boundary condition (member 0 is unperturbed) taken from a Gaussian distribution with a zero mean and a standard deviation set to 30 cm. This value was found empirically, in order to have a good spread at the boundary, which is then propagated to the variables computed by the model. As asserted, the perturbations are centred, having a null mean, and correlated in time. To do this, each perturbation at time t is obtained from a weighted average of a new perturbation and of the one at time $t-1$. This type of perturbations are classified as red noise and in the present case we used a decay time of two days."

R2.8 P9, L255: perhaps: "...towards the observations WHILE keeping the physical dynamics..."

Response: Corrected.

R2.9 p10, L298: I don't entirely see what is meant by "scalability". Can you please rephrase or clarify?

Response: We are referring to the computing scalability of the DA-EnSRF procedure on multiprocessor computers. The sentence has been rephrased as follow: "So in this case the simulations are 26535, but the computing scalability is high since the 61 simulations of the ensemble are independent and can be parallelized on multiple CPUs computers.

R2.10 P12, L351: These correlationS...

Response: Corrected.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-61, 2020.